doi:10.1006/jmbi.2000.4273 available online at http://www.idealibrary.com on IDEAL J. Mol. Biol. (2000) 304, 753–763

JMB

Functional Characterization of the Enhancer Blocking Element of the Sea Urchin Early Histone Gene Cluster Reveals Insulator Properties and Three Essential *cis*-acting Sequences

Raffaella Melfi[1]†, Franco Palla[2]†, Paola Di Simone[1] Claudia Alessandro[1], Larissa Calì[2], Letizia Anello[2] and Giovanni Spinelli[1]*

[1]*Dipartimento di Biologia Cellulare e dello Sviluppo (Alberto Monroy), Parco d'Orleans II, 90128, Palermo Italy*

[2]*Istituto di Biologia dello Sviluppo del Consiglio. Nazionale delle Ricerche, via Ugo La Malfa 153 90146, Palermo, Italy*

**Corresponding author*

Insulator elements can be functionally identified by their ability to shield promoters from regulators in a position-dependent manner or their ability to protect adjacent transgenes from position effects. We have previously reported the identification of a 265 bp *sns* DNA fragment at the 3′ end of the sea urchin H2A early histone gene that blocked expression of a reporter gene in transgenic embryos when placed between the enhancer and the promoter. Here we show that *sns* interferes with enhancer-promoter interaction in a directional manner. When *sns* is placed between the H2A modulator and the inducible *tet* operator, the modulator is barred from interaction with the basal promoter. However, the *tet* activator (tTA) can still activate the promoter, even in the presence of *sns*, demonstrating that *sns* does not interfere with activity of a downstream enhancer. In addition, the H2A modulator can still drive expression of a divergently oriented transcription unit, suggesting that *sns* does not inhibit binding of transcription factor(s) to the enhancer. To identify *cis*-acting sequence elements within *sns* which are responsible for insulator activity, we have performed *in vitro* DNase I footprinting and EMSA analysis, and *in vivo* functional assays by microinjection into sea urchin embryos. We have identified three binding sites for protein complexes: a palindrome, a direct repeat, and a C + T sequence that corresponds to seven GAGA motifs on the transcribed strand. Insulator function requires all three *cis*-acting elements. Based on these results, we conclude that *sns* displays properties similar to the best characterized insulators and suggest that directional blocking of enhancer-activated transcription by *sns* depends on the assembly of distinct DNA-protein complexes.

Keywords: histone genes; enhancer blocking; insulator; H2A enhancer; microinjection

Introduction

Proper temporal and spatial regulation of gene expression requires the orderly and efficient interaction of transcription factors with their cognate sites. The elucidation of how this might occur is one of the major challenges in molecular biology. If transcription units are organized into independent functional domains, enhancers could activate transcription from a promoter within the same domain but would be restrained from interacting with promoters in external domains. Insulators seem to be involved in the organization of the eukaryotic genome into domains of gene expression.[1–4] Insulators have been identified because they interfere with

†The first two authors contributed equally to this work.

Abbreviation used: EMSA, electrophoretic mobility shift assay.

E-mail address of the corresponding author: spinelli@unipa.it

0022-2836/00/050753–11 $35.00/0

enhancer-promoter communication only when positioned between them; for this reason they are generally called enhancer blocking elements. Although mechanisms by which these elements function in the normal context are not fully understood, it is thought that binding of chromosomal proteins to specific insulator sequences and subsequent protein-protein interactions can directionally restrict enhancer activity (for recent reviews see:[5,6]). The best characterized of these enhancer blocking elements are the *Drosphila*, scs, scs' and gypsy elements,[7–10] and the chicken HS4 insulator.[11,12] In addition to enhancer blocking activity, insulators can protect a transgene from chromosomal position effects when placed in flanking positions. However, only the chicken HS4[13,14] and the *Drosphila*, gypsy[15] elements seem to confer a barrier which prevents heterochromatin from spreading. Gypsy elements can also buffer a promoter from the silencing effect of the polycomb responsive element (PRE).[16] Elements which counteract silencing and act as heterochromatin boundaries have also been identified in yeast.[17–19] However, it is not known whether the yeast elements can also act as insulators of enhancer activity or display only chromatin boundary functions.[20]

Insulator DNA elements seem to be present in the tandemly repeated sea urchin early histone genes. The five early histone genes are expressed coordinately after meiotic maturation of the egg, and in early cleavage embryos until the blastula stage.[21] Although there is a single enhancer within the histone repeat unit, the 30 bp modulator of the H2A gene[22,23] each gene within the repeat is apparently regulated by gene-specific transcriptional elements.[24–26] From this observation, we hypothesized that there might be chromosomal elements which would direct and restrict H2A modulator function to its cognate H2A promoter. Subsequently, we identified a 265 bp sequence at the 3′ end of the H2A early histone gene (see map in Figure 1), that, in microinjected sea urchin embryos, showed a blocking activity when placed between the enhancer and the promoter. We termed this sequence silencing nucleoprotein structure (*sns*). Although proof of the directionality of enhancer blocking activity was lacking, experimental evidence suggested that *sns* is an insulator of enhancer function rather than a general repressor of transcription. *sns* blocked enhancer-promoter interaction only when interposed, in either orientation, between a multiple array of the H2A modulator/enhancer and the basal thymidine kinase (tk) promoter. No other position influenced transgene expression. Furthermore, *sns* interfered with enhancer function but not with the activity of the basal promoter, in that it mantained the capacity to silence transgene expression when it was placed at a distance of 2.7 kb from the promoter.[28] Interestingly, our results suggested evolutionary conservation of enhancer blocking mechanisms, as *sns* was able to shield promoters from viral enhancers in transient and stably transfected human cell lines, as well as in *Xenopus laevis* oocytes[28] (unpublished observations).

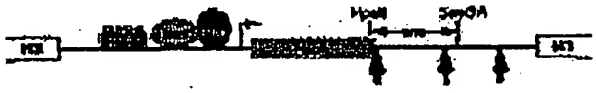

Figure 1. Map of the H2A transcription unit. CCAAT (CBF), modulator (MBF), putative GAGA factor binding sites in the H2A promoter, and the transcription start site are indicated. The *sns* fragment spanning from the *Hpa*II to the *Sau*3AI restriction sites starts 14 nucleotides upstream of the stop codon of the H2A gene and resides in a region where three micrococcal nuclease hypersensitive sites (vertical thick arrows) appear at gastrula stages.[27]

Here we have further characterized the properties of the sea urchin *sns* DNA fragment. We present evidence demonstrating that *sns*, as in the case of the best characterized insulators, affects only enhancers located distally from the promoter and displays a directionality in enhancer blocking activity. In addition, we report the identification of three protein binding sites within *sns* that, as demonstrated by functional assays in transgenic sea urchin embryos, are collectively required for insulator activity.

Results

sns interferes with enhancer function in a directional manner

Utilizing RNase protection assays, we previously demonstrated the ability of *sns* to inhibit enhancer-activated expression of a transgene in microinjected sea urchin embryos only when interposed between an array of the H2A modulator/enhancer and the *tk* promoter.[28] Because basal expression from *tk* or other viral promoters in sea urchin is negligible, effects on transcription were dramatic, with previously abundant transcript levels becoming undetectable. We have used the same approach to investigate whether *sns* evinces other behaviors expected of insulators. We placed *sns* between two enhancers in a chloramphenicol-acetyl transferase (CAT) gene transcription unit (Figure 2(a), construct 2). As a distal enhancer, we used the sea urchin H2A modulator array, containing at least four binding sites for the MBF-1 *trans*-activator; the proximal enhancer was the heptamerized tetracycline (*tet*) operator. In HeLa cells, the *tet* operator is induced upon binding of the Tet repressor-VP16 activation domain chimera (tTA *trans*-activator) and stimulates transcription of a reporter gene by several orders of magnitude.[29] The *tTA* gene was placed under the control of the multiple modulator elements and the *tk* minimal promoter (construct A). This construct was coinjected with the *CAT* reporter gene driven by the *tet*

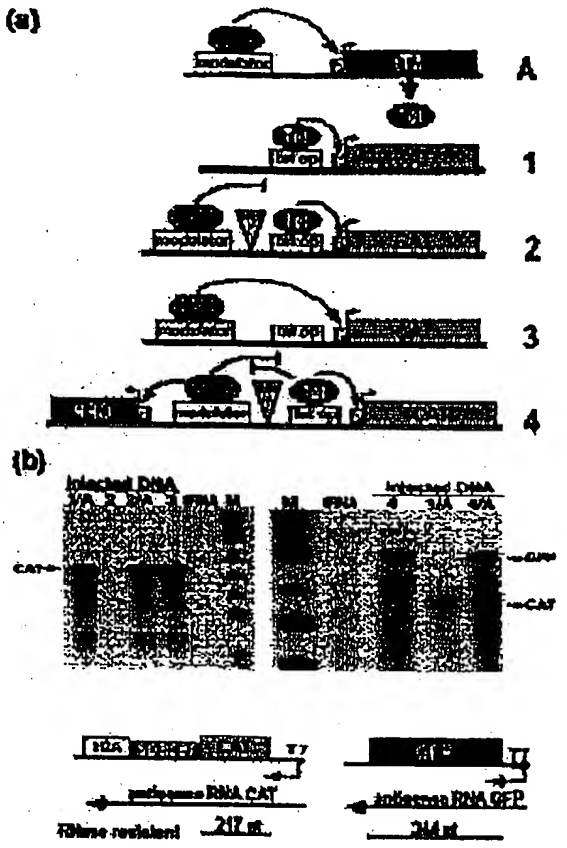

Figure 2. Polar and directional effects of *sns* on enhancer-promoter interaction. (a) Schematic representation of the microinjected plasmids. Binding of the MBF and the tTA transactivators, respectively, to the modulator and *tet* operator arrays is also indicated. Curved arrows refer to the activation of promoter by the bound factor; small arrows point to the transcription start site. (b) Total RNAs from microinjected embryos were processed to detect enhancer-activated transgene expression by RNase protection using antisense ^{32}P-labelled probes transcribed *in vitro* from the constructs illustrated below the gel images. Arrows point to the 217 nucleotide CAT and 344 nucleotide GFP RNase-resistant fragments. tRNA was used as negative RNA control. The RNase digestion products and end labelled *Hpa*II-digested pBluescript DNA (M) were run on denaturing polyacrylamide gels. Coinjection of constructs 1 and A (lanes 1/A) leads to CAT expression. In the absence of tTA, construct 2 is silent because the modulator is blocked by *sns* (lane 2). Coinjection of constructs 2 and A (lane 2/A) *trans*-activates *CAT* gene expression. Construct 3 expresses the *CAT* transgene because the *tet* operator does not interfere with the modulator (lane 3). Construct 4, in the absence of tTA, expresses only the GFP transgene (lane 4); if coinjected with A both the *CAT* and the *GFP* reporter genes are *trans*-activated (lane 4/A). Hence, *sns* does not block the binding of the transcription factors to the enhancers.

operator (construct 1). We predicted that expression of the tTA would elicit *trans*-activation of the transgene. As shown in Figure 2(b) (lane 1/A) this was indeed the case. Next, we tested the constructs with two enhancers. As expected, construct 2 containing *sns* between the MBF-1 and tTA binding sites was transcriptionally silent (lane 2), indicating that *sns* blocked the *trans*-activating function of the MBF-1 and that the *tet* operator was inactive in the absence of tTA. When the activator expression plasmid (construct A) was coinjected with construct 2, *trans*-activation of the transgene occurred (lane 2/A). The intensity of the CAT mRNA band detected in embryos injected with the two different plasmid combinations was almost identical (compare lane 1/A and 2/A), suggesting that neither the modulator nor *sns* sequences affected the extent of activation by tTA. In addition, the *tet* operator sequences did not interfere with the enhancer activity of the modulator, as similar levels of transgene expression were detected in embryos injected with construct 3 (lane 3). In summary, these experiments strongly suggest that *sns*, like chromatin insulators, has the ability to block the distal enhancer from communicating with the promoter but has no influence on the proximal one, when situated between the two.

We also investigated *sns* behavior in the context of a bidirectional transcription construct. A construct was made in which the modulator array and *tet* operator direct expression of two divergently transcribed reporter genes, encoding either CAT or green fluorescent protein (GFP). The *sns* sequence was inserted between the modulator and *tet* operator (construct 4 in Figure 2(a)). RNase protection assays were performed with RNA extracted from transgenic embryos, utilizing probes for both CAT and GFP in the same hybridization reaction. Only *CAT* mRNA was detected in embryos microinjected with constructs 1 and A (Figure 2(b), right panel, lane 1/A). As expected, in the absence of the tTA activator, the *CAT* transgene was not expressed in embryos injected with the bidirectional transcription unit (lane 4), presumably because *sns* interrupted the interactions between MBF-1 and the basal transcription apparatus. However, *sns* did not restrain MBF-1 from activating the divergent *GFP* transcription unit (lane 4). Subsequently, expression of tTA allowed for *trans*-activation of the *CAT* transcription unit (lane 4/A). From these results we conclude that *sns* blocks enhancer activity in a directional manner.

In vitro binding of proteins to *sns* sequences

To identify nuclear protein binding sites within *sns*, we performed DNase I footprint analysis and electrophoretic mobility shift assays (EMSA) with nuclear extracts from gastrula stage embryos. Two DNase I protected regions, defined as Box A and Box B, were mapped to both strands in the 5′ half of *sns* (Figure 3). The specificity of protein-DNA interaction was assessed by oligonucleotide compe-

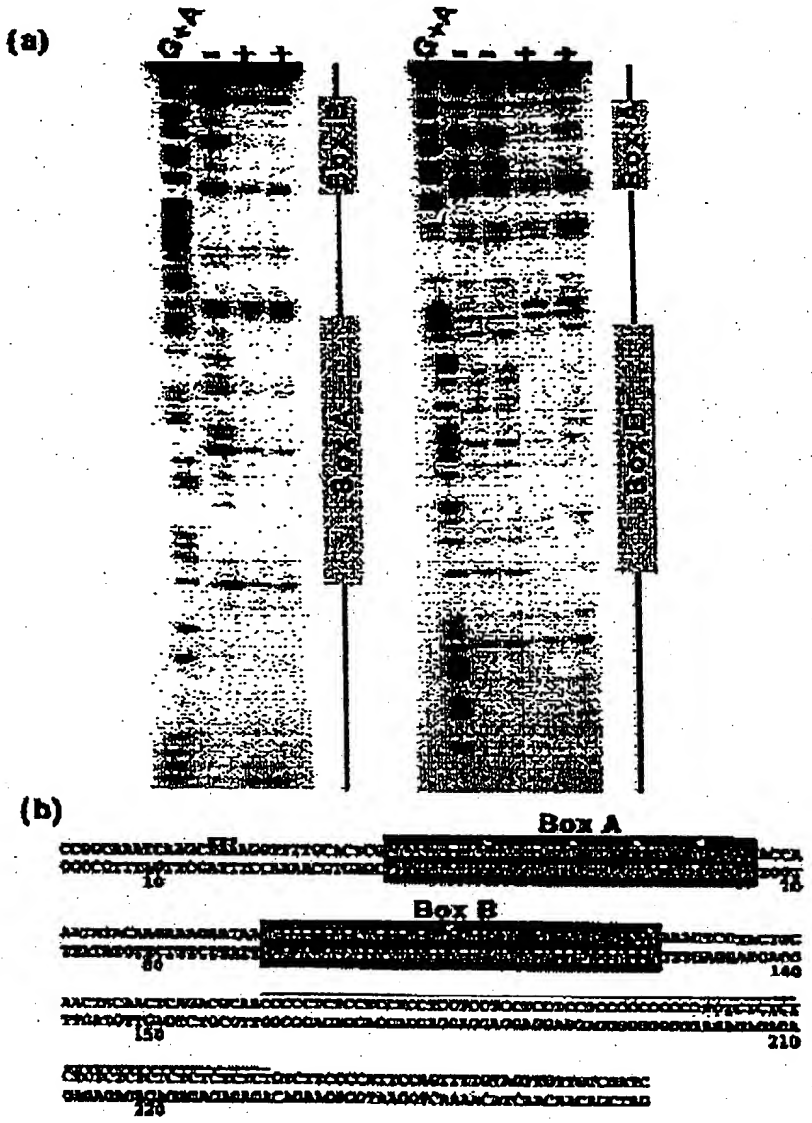

Figure 3. DNase I footprinting of an end-labelled *sns* sub-fragment spanning from nucleotides 2 to 157. (a) Sense strand (left) and antisense strand (right) were incubated with either BSA (lanes-) or nuclear extracts from gastrula stage embryos and digested with 5 μg of DNase I for three to five minutes on ice (lanes+). Digestion products were analyzed together with the cleavage products of the G + A sequence reaction on denaturing polyacrylamide gels. (b) Nucleotide sequence of the *sns* element. Asterisks mark the stop codon of the H2A gene. Sequences of protected Box A and Box B are in black boxes. Arrows inside boxes point to the Box A inverted and direct repeats and to the Box B direct repeat. Pyrimidine stretches are overlined.

tition experiments in EMSA. Figure 4 shows that both DNA-protein complexes were suppressed by an excess of unlabelled homologous probe, while they were not affected by an excess of unlabelled heterologous sequences. As indicated in the sequence shown in Figure 3(b) and in the drawing of Figure 4, Box A contains two notable sequence features: a C + A perfect direct repeat (DR), and immediately downstream the palindrome (IR) which is one of the *cis*-acting elements involved in 3′ RNA processing.[30] Because the IR sequence alone competed as efficiently as the entire Box A, we conclude that the palindrome is the protein binding site within Box A.

To search for further protein binding sites, we analyzed the pyrimidine stretch (C + T) at the 3′ end of *sns* sequences. This fragment contains 14 TC repeats that in the bottom strand correspond to seven GAGA sequences. EMSA analysis with nuclear extracts demonstrated that the C + T rich fragment formed a predominant DNA-protein complex that was specifically competed by an

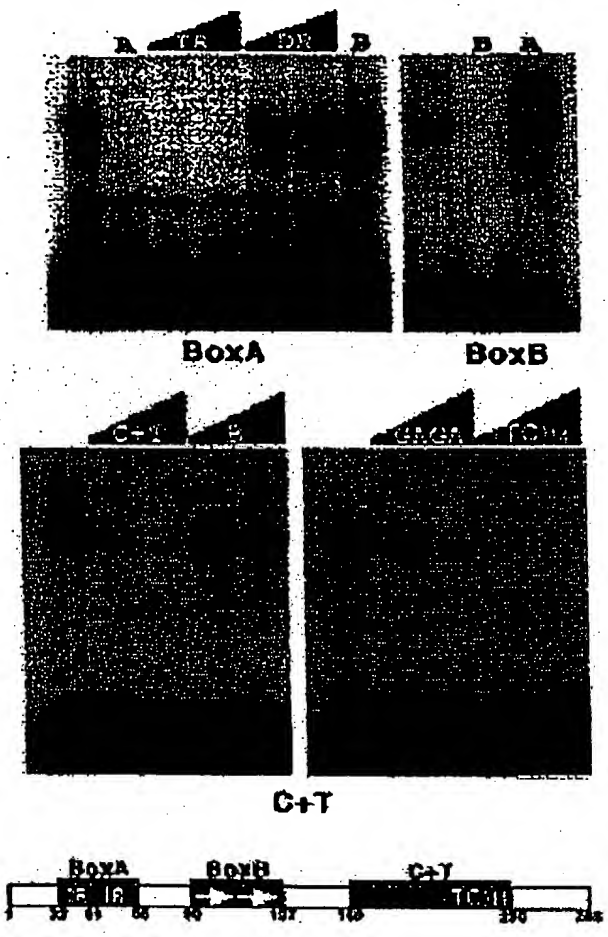

Figure 4. EMSA analysis of nuclear protein binding sites within *sns*. The three end labelled probes Box A, Box B and C+T, are underlined in the schematic drawing of the *sns* fragment. DR and IR refer, respectively, to the direct repeat and palindrome of Box A; white arrows in the Box B indicate a tandem repetition; (TC)14 refers to the 14 repetitions of the TC doublet. The C+T fragment was obtained by PCR amplification. All other probes were obtained by annealing complementary oligonucleotides. In competition experiments, nuclear extracts were pre-incubated with an excess of unlabelled homologous or heterologous probes prior to the addition of 1 ng of the labelled probe. The amounts used were: 100 ng for Box A (A) and Box B (B); 50 ng and 100 ng for IR, DR, C+T, GAGA (the GA repeats located upstream the H2A modulator), and (TC)14. The DNA-protein complexes were resolved by polyacrylamide gel electrophoresis.

excess of the homologous fragment (Figure 4). Of particular significance, protein binding was also specifically competed when nuclear extracts were pre-incubated either with an excess of a sequence containing four GAGA repeats, which is located upstream of the *H2A* modulator (see Figure 1), or with an oligonucleotide containing the 14 TC dinucleotides found at the 3' end of the pyrimidine region (see sequence in Figure 3). The former competition was slightly less efficient, perhaps due to the presence of fewer (eight) TC dinucleotide repeats. These observations demonstrate the binding of nuclear protein(s) to GAGA sequences in sea urchin and suggest that a putative GAGA factor might contribute to the enhancer blocking function of *sns*.

Deletion of either the Box A palindrome or the 3' CT repeats abolishes *sns* insulator function

We used the enhancer blocking assay to test the effect of 5' and 3' deletions of the *sns* fragment on the expression of a transgene driven by the *H2A* modulator in transgenic sea urchin embryos. The *sns* deletion mutants shown in Figure 5(a) were cloned between multiple copies of the 30 bp modulator/enhancer of the H2A histone gene and the tk promoter of the M30-CAT reporter plasmid (Figure 5(b)). Resulting constructs were microinjected into sea urchin eggs, embryos raised till gastrula stages and processed to determine *CAT* transgene expression by RNase protection analysis. Results depicted in Figure 5(c) are representative of several microinjection experiments. In agreement with our previous reports,[23,31] in the presence of one or several copies of the 30 bp histone H2A modulator sequence, transcriptional activation from the tk promoter occurs efficiently, as evidenced by abundant transgene transcripts (Figure 5(c), lanes 3, 9, 11). These M30-CAT constructs demonstrated once again the enhancer blocking function of the intact *sns* (lane 5). Deletions from either the 5' or the 3' that remove Box A(Δ sns), or the pyrimidine rich sequence (ΔIII sns), respectively, impaired the blocking activity of *sns* (lanes 4 and 13). In fact, levels of CAT transcripts were comparable to the construct lacking *sns* (lanes 3 and 11). As expected, 5' deletions that

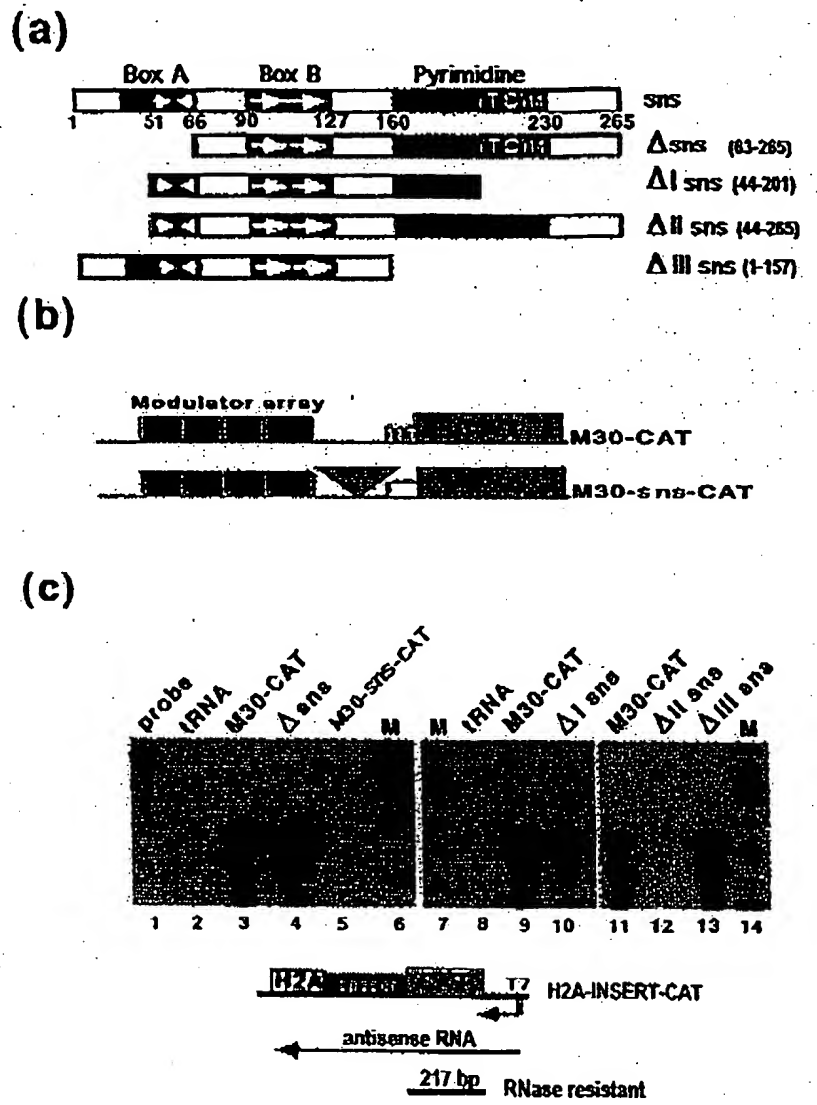

Figure 5. Functional activity of the *sns* deletion fragments. (a) Maps of *sns* and of the different deletion fragments assayed for enhancer blocking activity. (b) Schematic drawing of the microinjected plasmids. *sns* or the *sns* deletion fragments were inserted between the H2A modulator array and the tk promoter of M30-CAT to generate the M30-sns-CAT plasmids. (c) A ^{32}P-labelled antisense CAT RNA (lane 1), transcribed *in vitro* from the H2A-INSERT-CAT was used to perform RNase protection assays on microinjected embryos at gastrula stage. Electrophoretic analysis of the RNase digestion products was carried out on denaturing polyacrylamide gels. Lanes 3, 9 and 11: microinjection of the positive control M30-CAT to monitor enhancer-activated expression of the transgene. Lane 5: microinjection of M30-sns-CAT to monitor enhancer blocking activity of *sns*. Lanes 4, 10, 12 and 13: microinjection of reporter plasmids carrying, respectively, Δsns, ΔI sns, ΔII sns and ΔIII sns; only ΔII *sns* maintained the ability to attenuate the enhancer. Lanes 2 and 8: tRNA negative control. Lanes 6, 7 and 14: pBluescript *Hpa*II-digested end-labelled DNA markers.

left the palindrome intact, (ΔII sns) exhibited wild-type *sns* enhancer blocking activity (lane 12). Finally, removal of the TC repeats (ΔI sns) from ΔII *sns*, abolished the ability of *sns* to affect enhancer-promoter interaction (lane 10). Altogether, these results are consistent with the nuclear protein binding sites defined above and indicate that the Box A palindrome and the GAGA sites are essential for *sns* to block communication between the modulator and the tk promoter.

Box B is also essential for enhancer blocking activity

The experiments described in the previous sections suggest that the enhancer blocking function of *sns* relies on the assembly of protein complexes

at the Box A palindrome and at the GAGA sites. Because binding of proteins to Box B was also detected, we investigated whether these interactions were also essential for *sns* activity. Toward this end, we performed an *in vivo* competition experiment. We have previously used this approach to demonstrate that binding of the MBF-1 transcription factor to the modulator is required for activation of a transgene driven by the histone H2A promoter.[31] As indicated in Figure 6, sea urchin embryos were injected with the *sns*-containing transgene construct together with increasing amounts of ligated oligonucleotides containing either Box B (lanes 3, 4) or the Box A (lane 6) sequences. As levels of enhancer-activated transgene transcripts were similar to those seen with M30-CAT plasmid (lane 5), these results demonstrate that either oligonucleotide prevented enhancer blocking (lanes 2,7). Hence, titration of either Box A or Box B binding proteins by injecting their target sites impaired the ability of *sns* to block enhancer-promoter interaction.

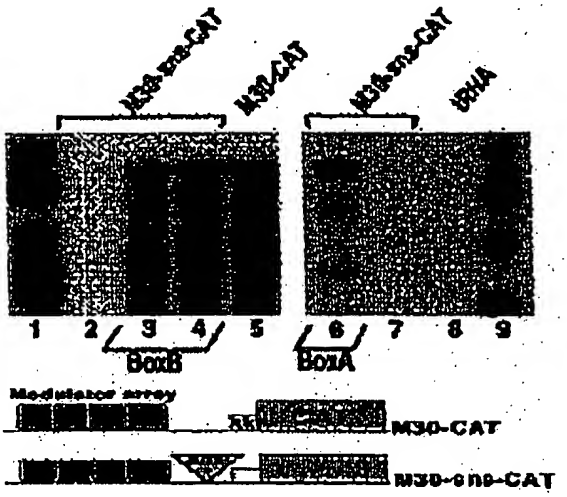

Figure 6. *In vivo* competition of *sns* function. Transgenic embryos were obtained by microinjecting the constructs drawn below the fluorograph with or without an excess of either BoxB or BoxA oligonucleotides. Black boxes represent the modulator array, the large shaded triangle the *sns* fragment. RNase protection experiments were carried out with total RNA as described in the legend to Figure 5. Lanes 2 and 7: injection of M30-sns-CAT; the enhancer is blocked. Lanes 3, 4 and 6: co-injection of M30-sns-CAT and 40-fold (lane 3) or 70-fold (lane 4) excess of ligated BoxB oligonucleotide or 70-fold (lane 6) excess of ligated BoxA oligonucleotide relieves the block. Lane 5: injection of M30-CAT; enhancer activity of the modulator array. Lanes 1 and 9 show relevant bands of DNA markers.

Discussion

Insulators are a new class of genetic elements that can modulate the activity of enhancer or other regulative sequences.[3,5] The few elements identified principally in *Drosophila* and chicken display two important characteristics: polarity and directionality of the effects of insulation of enhancer activity.[1,2] The former signifies that only enhancers located distally from the promoter with respect to the site of insertion of the insulator are attenuated in the interaction with the promoter. The second feature is that insulators do not prevent a blocked enhancer from activating transcription from a divergent promoter.[32,33] Consistently, we have shown that *sns* when placed between two enhancers, insulated the promoter-distal modulator without affecting the function of the downstream *tet* operator. In addition, *sns* did not interfere with the *trans*-activating capacity of the modulator in the other direction. Taken together, these results rule out that insertion of *sns* between enhancer and promoter represses enhancer-promoter interaction by enhancer inactivation, for example by inducing local assembly of a repressive chromatin structure.

As first shown in *Drosophila*, the directional enhancer blocking activity of insulator elements depends on the assembly of specific DNA-protein complexes. The gypsy insulator is perhaps the best-studied system with respect to the characterization of protein components that interact with insulator DNA. One of these components, the suppressor of Hairy-wing [su(Hw)] protein, binds to a reiterated target sequence[34] and recruits the second component, the mod(mdg4) protein[35] that displays properties characteristic of trithorax-group (trxG) genes.[36] The BEAF protein binds to the *scs'* insulator[37] which characterizes a class of chromosomal elements found at many loci.[38] Interestingly, seven tandem copies of an oligonucleotide containing BEAF binding sites has partial enhancer blocking activity.[37] The capability of the chicken HS4 insulator to interfere with enhancer-promoter interaction resides in a 42 bp fragment that contains a binding site for the CTCF transcription factor.[39] Binding of CTCF occurs also to several vertebrate insulators and to the unmethylated ICR (imprinting-control region) that displays enhancer blocking activity to control imprinted expression of the Igf2 gene.[40,41] Therefore, it is not surprising that the directional enhancer blocking activity of *sns* depends on specific DNA-protein interactions. It is of some interest that, while the enhancer blocking capacity of the *gypsy*, *HS4*, and to some extent the *scs'*, insulators relies on the recognition of a single or a reiterated binding site, *sns* contains three different *cis*-acting elements. Our results strongly suggest that all of these are needed to prevent enhancer-promoter interaction. In fact, deletion of either the Box A palindrome or the 14 TC repeats completely impaired *sns* function. Furthermore, microinjection of excess Box A or Box B and very recently GAGA (not shown) binding sites relieved the inhibition of the modulator in the *sns*-containing constructs. The most plausible explanation of the *in vivo* competition results is that the excess of

binding sites titrated, either directly or indirectly, the factors responsible for the enhancer blocking activity. Based on these observations, we speculate that *sns* achieves directional enhancer blocking activity by cooperative interactions between all three different DNA binding proteins or protein complexes.

Our results demonstrate that, within Box A, only the palindrome is required for enhancer blocking activity. Deletion of the 5' most direct repeats, upstream of the palindrome, does not impair *sns* function. In agreement with this observation, oligonucleotides containing the direct repeats failed to compete for binding of factors to Box A and did not form specific protein-DNA complexes (not shown). Of some interest, the palindrome forms a stem-loop RNA structure, highly conserved among the non-polyadenylated histone mRNAs, from sea urchin, and *Drosophila* to mammals and represents one of the signals recognized by 3' pre-histone mRNA processing machinery.[30]

A second *cis*-acting element was identified within the pyrimidine tract that contains seven GAGA repeats in the inverted orientation. Based upon EMSA analysis, specific protein interactions occur at the GAGA sites of *sns* and presumably at GAGA sites located upstream of the H2A modulator. Because the enhancer blocking activity of *sns* is independent of orientation,[28] it is reasonable to assume that protein(s) related to a *Drosophila* factor which binds GAGA sequences might be involved in the mechanism that interrupts the interaction between enhancer and promoter in sea urchin. *Drosophila* GAGA factor is a DNA binding protein involved in chromatin remodelling processes.[42] GAGA factor alleviates, in combination with NURF, the repressive effect of chromatin[43] and participates in the assembly of the silencing Polycomb proteins at PRE.[44] Interestingly, binding of factors to GAGA sites occurs in the spacer between the *Drosophila* H3 and H4 histone genes,[45] and recent evidence indicates a direct involvement of GAGA factor in insulator activity. GAGA factor binding sites, found at the PRE adjacent to the Fab-7 insulator, cooperate with Fab-7 to maintain the specific parasegment domain of expression of the Abdominal-B gene.[46,47] In addition, mutation of GAGA sequences within the insulator of the even-skipped locus affects directional blocking of the iab-5 enhancer.[48] Despite the similarity of the binding site and the apparent involvement in insulator function, the sea urchin protein differs from the *Drosophila* GAGA factor because a *Drosophila* polyclonal anti-GAGA factor antibody failed to supershift the *in vitro* assembled nuclear protein-DNA complex from sea urchin (not shown). The cloning of the sea urchin GAGA factor encoding gene should clarify whether the *Drosophila* and sea urchin factors are evolutionary and functionally related. One working hypothesis, currently under investigation, is that interactions between the proteins of *sns* and the proteins bound to the GAGA sites of the H2A promoter, prevent the H2A enhancer from acting promiscuously to activate transcription of heterologous early histone promoters.

With the exception of the GAGA element, the *sns* insulator sequence motifs are distinct from those described for other insulators. However, there is some evidence to suggest that these insulator sequences and their binding factors are also evolutionarily conserved. Very similar sequences are present in equivalent positions in the histone H2A transcription unit of the sea urchin *Psammechinus miliaris* (not shown). In addition, we have recently found that *sns* can insulate a viral enhancer upon stable integration in human chromatin (unpublished) and that at least two of the identified *cis*-acting insulator sequence elements, Box B and TC dinucleotide repeats (Box A did not show DNA binding activity in our conditions), interact specifically with human nuclear proteins of two different cell types (unpublished results).

In conclusion, we have extended our previous characterization of the *sns* element by the demonstration that *sns* acts equivalently to previously well-characterized insulators in a number of ways. We have now identified *cis*-acting sequences required for directional enhancer blocking activity, which may be evolutionarily conserved, and include novel sequences. Our studies have significant implications both for the control of early histone gene regulation in sea urchins, and for more general mechanisms of insulator action. In addition, these sequences may prove to have practical applicability in genetic engineering situations where insulator action might be beneficial.

Materials and Methods

Construction of plasmids

Plasmids, schematically drawn in Figure 2, were obtained as follows. Plasmid A that expresses the tTA activator, was constructed by the substitution of the CMV promoter of the pUHD 15.1 vector[29] with a fragment containing an array of the modulator sequences and the *tk* promoter. The pUHD 15.1 plasmid was digested with *Xho*I and *Xba*I simultaneously, filled in and ligated with a blunt ended DNA fragment containing the modulator sequences. Plasmid 1 that expresses the *CAT* gene under the control of the *tet* operator and the CMV promoter, was constructed by cloning the *tet* operator and the CMV promoter from the pUHD 10.3 plasmid[49] into the *Xho*I restriction site of the pBL-CAT3 vector.[50] Plasmid 3 that expresses the *CAT* gene under the control of two enhancers, the *tet* operator and the modulator, was obtained by cloning the 180 bp *Hin*dIII-*Xba*I DNA fragment containing the modulator repeats, derived from M30-CAT, in the *Hin*dIII-*Xba*I-digested plasmid 1. In construct 2, the *Hin*dIII-*Xba*I DNA fragment containing *sns* was cloned into plasmid 3 between *tet* operator and the modulator sequences. The EGFP gene (Clontech) fused to the *tk* promoter was cloned in inverted orientation upstream the modulator of plasmid 2, to generate the construct containing the two divergent transcription units (plasmid 4). To generate the plasmid M30-CAT an array of the H2A modulator/enhancer sequences was cloned into the *Sal*I site upstream of the

tk promoter of the pBL-CAT3 vector. To obtain the plasmid M30-sns-CAT, *sns* was inserted into the *Xba*I site of M30-CAT, between the modulator and the *tk* promoter. The *sns* fragment was isolated by *Hpa*II and *Sau*3A restriction enzyme digestion of the *Sac*I subclone of the PH70 histone DNA.[21] The Δsns, ΔI sns, ΔII sns and ΔIII sns deletion fragments were obtained from *sns* by restriction enzyme digestion or PCR amplification. These fragments were cloned respectively in the *Xba*I and *Bam*HI restriction sites located in the polylinker between the modulator and the tk promoter of the M30-CAT. The plasmids used for the RNase protection experiments, H2A-INSERT-CAT and GFP, were obtained as follows. The former contains a 258 bp pUC fragment inserted between the H2A promoter and a CAT coding subregion as described.[23] The latter bears a 344 bp fragment of the EGFP gene.

DNase I protection and electrophoretic mobility shift assays

Sense and antisense *sns* end-labelled fragments, spanning the region between nucleotides 2 and 157, were obtained by PCR. BSA or nuclear extracts from *P. lividus* embryos, prepared as described,[51] were preincubated with 10 μg of poly(dA-dT)·(dA-dT) in 50 μl of 10 mM Hepes (pH 7.9), 60 mM KCl, 1 mM DTT, 1 mM EDTA, 8% (v/v) glycerol for five minutes before the addition of the labelled fragments. After 20 minutes incubation, samples were digested with 5 μg of DNase I in 2.5 mM $MgCl_2$ for three to five minutes at 4°C, phenol extracted and loaded on 8% (w/v) denaturing polyacrylamide gels. In EMSA experiments, three different *sns* probes were assayed for binding activity *in vitro*: Box A, Box B double-stranded oligonucleotides (see next paragraph for sequences) and the C + T containing subfragment (spanning nucleotides 137 and 262) obtained by PCR. 1 ng of end labelled Box A , Box B and C + T probes were incubated, for 30 minutes on ice, with 5 μg of nuclear extracts from sea urchin embryos at gastrula stage and 2 μg of poly(dI-dC)·(dI-dC) in 20 μl of 10 mM Hepes (pH 7.9), 60 mM KCl, 1 mM DTT, 1 mM EDTA, 4% Ficoll. In the competition experiments, unlabelled homologous or heterologous probes in the amounts described in the legend to Figure 4 were added to the preincubation mixture prior to the addition of the extracts. Binding reactions were run on non-denaturing polyacrylamide gels in 50 mM Tris, 50 mM H_3BO_3 and 2 mM EDTA (pH 8.3).

List of oligonucleotides

Only the sense strand is reported. Residues of the following oligos refer to the first nucleotide of the *sns* sequence shown in Figure 3 (accession number Y09062): Box A: CAAACCTCAACACCTCAACGGCCCTTATCAGGGCCACCA (spanning nucleotides 32 to 70); Box B: AGTCTCTGTAATTCATAATAGTCTCTGTAATTCATAAT (spanning nucleotides 90 to 127); DR: CAAACCTCAACACCTCAAC (spanning nucleotides 32 to 51); IR: CTCAACGGCCCTTATCAGGGCCACCAA (spanning nucleotides 45 to 71); (TC)14:CCTCTCTCTCTCTCTCT CTCTCTCTCTCTCTGTCT (spanning nucleotides 200 to 234). The GAGA oligonucleotide, GGGAGGAGAGAGAGAGAGAGA, spans nucleotides −200 to −177 from the transcription starting site of the H2A promoter.

Microinjection and transgene expression

Plasmids were linearized with *Hin*dIII or *Cla*I restriction enzymes whose sites are located in the polylinker regions, 5′ and 3′ to the CAT sequence, respectively. Linearized plasmids were brought to a total concentration of 240 μg/ml with threefold excess of *Hin*dIII-digested *P. lividus* sperm DNA of roughly 5 kb. Unfertilized eggs from a mature female were dejellied in pH 5.0 sea water, immobilized on Petri dishes with 1% (w/v) protamine sulfate and injected into the cytoplasm with 1-2 pl DNA in 25% glycerol solution. Microinjected eggs were fertilized and embryos raised up to the early gastrula stage (14 hours development). 3000 uninjected embryos at the same developmental stage were then added as carrier, and embryos collected by low speed centrifugation. Total RNA was extracted by one hour incubation at 55°C in a solution containing 200 μg/ml Proteinase K and phenol extraction. Nucleic acid samples were digested with RNase-free DNase I, and RNA from 50 microinjected embryos was hybridized with [^{32}P]UTP-labelled antisense CAT or EGFP RNA transcribed *in vitro* from the H2A-spacer-CAT and GFP plasmids. Hybridisation conditions, RNase digestion and gel fractionation of the RNase-resistant hybrids were as previously described.[52]

Acknowledgments

This work was supported by grants of MURST ("Ministero dell' Università e della Ricerca Scientifica e Tecnologica", 60% and Cofinanziamenti Programmi di Ricerca Scientifica di Interesse Nazionale, 99). Many thanks are due to Dr Sylvia Evans for reading and editing the manuscript and to Dr Bujard and co-workers for DNA clones containing *tet* operators and the *tTA* activator gene.

References

1. Gerasimova, T. I. & Corces, V. G. (1996). Boundary and insulator elements in chromosomes. *Curr. Opin. Genet. Dev.* 6, 185-192.
2. Geyer, P. (1997). The role of insulator elements in defining domains of gene expression. *Curr. Opin. Genet. Dev.* 7, 242-248.
3. Kellum, R. & Elgin, S. (1998). Punctuating the genome. *Curr. Biol.* 8, R521-R524.
4. Udvardy, A. (1999). Dividing the empire: boundary chromatin elements delimit the territory of enhancers. *EMBO J.* 18, 1-8.
5. Bell, A. C. & Felsenfeld, G. (1999). Stopped at the border, boundaries and insulators. *Curr. Opin. Genet. Dev.* 9, 191-198.
6. Dorsett., D. (1999). Distant liaisons: long-range enhancer-promoter interactions in *Drosophila*. *Curr. Opin. Genet. Dev.* 9, 505-514.
7. Kellum, R. & Schedl, P. (1991). A position-effect assay for boundaries of higher order chromosomal domains. *Cell*, 64, 941-950.
8. Kellum, R. & Schedl, P. (1992). A group of scs elements function as domain boundaries in an enhancer-blocking assay. *Mol. Cell. Biol.* 12, 2424-2431.

NURF in transcriptional activation of chromatin. *Mol. Cell*, 1, 141-150.
44. Horard, B., Tatout, C., Poux, S. & Pirrota, V. (2000). Structure of a polycomb response element and *in vitro* binding of polycomb group complexes containing GAGA factor. *Mol. Cell. Biol.* 20, 3187-3197.
45. Gilmour, D. S., Thomas, G. H. & Elgin, S. C. R. (1989). *Drosophila* nuclear proteins bind to regions of alternating C and T residues in gene promoters. *Science*, 245, 1487-1490.
46. Mihaly, J., Hogga, I., Gausz, J., Gyurkovics, H. & Karch, F. (1997). *In situ* dissection of the Fab-7 region of the bithorax complex into a chromatin domain boundary and a Polycomb-response element. *Development*, 124, 1809-1820.
47. Hagstrom, K., Muller, M. & Schedl, P. (1997). A Polycomb and Gaga dependent silencer adjoins the Fab-7 boundary in the *Drosophila* bithorax complex. *Genetics*, 146, 1365-1380.
48. Ohtsuki, S. & Levine, M. (1998). GAGA mediates the enhancer blocking activity of the eve promoter in the *Drosophila* embryo. *Genes Dev.* 12, 3325-3330.
49. Gossen, M., Freundlieb, S., Bender, G., Muller, G., Hillen, W. & Bujard, H. (1995). Transcriptional activation by tetracyclines in mammalian cells. *Science*, 268, 1766-1769.
50. Luckow, B. & Schütz, G. (1987). CAT construction with multiple unique restriction sites for the functional analysis of eukaryotic promoters and regulatory elements. *Nucl. Acids Res.* 15, 5490.
51. Palla, F., Bonura, C., Anello, L., Casano, C., Ciaccio, M. & Spinelli, G. (1993). "Sea urchin early H2A modulator binding factor 1 is a positive transcription factor also for the early histone H3 gene". *Proc. Natl Acad. Sci. USA*, 90, 6854-6858.
52. Di Bernardo, M., Russo, R., Oliveri, P., Melfi, R. & Spinelli, G. (1995). Homeobox-containg gene transiently expressed in a spatially restricted pattern in the early sea urchin embryo. *Proc. Natl Acad. Sci. USA*, 92, 8180-8184.

Edited by M. Yaniv

(Received 21 July 2000; received in revised form 10 October 2000; accepted 24 October 2000)